Math 217 Definitions and Main Theorems in order by Section number in the book

Professor Karen E Smith[1]

These are the definitions you must understand and be able to state *precisely* to succeed in Math 217. Some are different than in the textbook, which does not treat all concepts in the generality you will need. We also list the crucially important **Theorems** and **Propositions** you must know, and provide model proofs. In doing proofs in Math 217, you should refer to this document for the correct definitions you will need, and for an example of how to clearly write a proof.

This document **does not replace the textbook**, which also contains material you must know. In particular, the texbook contains many important examples and computational techniques that are crucial to succeeding in Math 217. For your convenience, we also provide a list of the vocabulary and techniques from the book you should be sure to understand, but we only list them briefly. **Please refer to the book to make sure you fully understand all the listed words and techniques.** The book's **Summaries** at the end of each section are especially useful.

1. Chapter 1

The main idea is systems of linear equations, and matrices as a tool to solve them.

Section 1.2. The book mentions vector spaces in 1.2 but does not give a careful definition until Chapter 4. A vector space is a **set with extra structure** whose elements will be called **vectors.** The extra structure consists of a natural **addition** and **scalar multiplication,** which must obey certain familiar axioms. Here is the precise definition:

Definition 1.1. A vector space is a set V, equipped with a rule for **addition** of any two vectors and for **scalar multiplication** of a vector by a scalar. The addition $+$ must satisfy the following axioms

(1) The set V is closed under addition: For any two vectors $\vec{v}$ and $\vec{w}$ of V, the sum $\vec{v}+\vec{w}$ is also in V.
(2) Addition is commutative: For all $\vec{v},\vec{w} \in V$, $\vec{v}+\vec{w} = \vec{w}+\vec{v}$.
(3) Addition is associative: For all $\vec{v},\vec{w},\vec{y} \in V$, $(\vec{v}+\vec{w})+\vec{y} = \vec{v}+(\vec{w}+\vec{y})$.
(4) There is an additive identity: that is, there exists $\vec{0} \in V$ such that $\vec{v}+\vec{0} = \vec{0}+\vec{v} = \vec{v}$ for all $\vec{v} \in V$.
(5) Every element has an additive inverse: for every $\vec{v} \in V$, there exists a vector $\vec{y} \in V$ such that $\vec{v}+\vec{y} = \vec{y}+\vec{v} = 0$.

[1]Thanks to Katie Maddon, Jordan Katz, Dominic Russell, Mark Spencer, Aarush Garg, Zhengxi Tan, Shrikar Thodlar, Miranda Riggs, Suki Dasher, Margaret Ehinger, Betty Qi for useful remarks.

The scalar multiplication must satisfy the following axioms

(6) The set V is closed under scalar multiplication: For any vector $\vec{v}$ in V and any scalar $\lambda \in \mathbb{R}$, the scalar multiple $\lambda\vec{v}$ is also in V.
(7) For two scalars $a, b \in \mathbb{R}$, we have $a(b\vec{v}) = (ab)\vec{v}$ for all vectors $\vec{v} \in V$.
(8) For $0 \in \mathbb{R}$, we have $0\vec{v} = \vec{0}$ for all $\vec{v} \in V$.
(9) For $1 \in \mathbb{R}$, we have $1\vec{v} = \vec{v}$ for all $\vec{v} \in V$.

And finally, scalar multiplication distributes over addition:

(10) $\lambda(\vec{v} + \vec{w}) = \lambda(\vec{v}) + \lambda(\vec{w})$ for all $\vec{v}, \vec{w} \in V$ and all $\lambda \in \mathbb{R}$.
(11) $(a + b)\vec{v} = a\vec{v} + b\vec{v}$ for all vectors $\vec{v} \in V$ and all scalars $a, b, \in \mathbb{R}$.

The only example the book gives at this point is the space $\mathbb{R}^n$ of column vectors (of size n) with the usual addition and scalar multiplication. We will call this the **coordinate space** of dimension n. Admittedly, this is a very important example, but there are many others.

Example 1.2. Examples of vector spaces:

(1) In multivariable calculus and physics, you learned a vector is a "directed magnitude" represented by an arrow. The set of all such vectors (say, in 3-space) forms a vector space with the usual notion of vector addition (placing arrows "head-to-tail") and scalar multiplication (scaling the magnitude). We call this vector space the "coordinate-free space" of dimension three, and denote it $\mathbb{E}^3$.
(2) Let $\mathcal{F}$ be the set of all functions from $\mathbb{R}$ to $\mathbb{R}$. Since high school, you have added such functions (by simply adding the outputs) and multiplied them by scalars $cf(x)$.The set $\mathcal{F}$ forms a vector space with these familiar notions of addition and scalar multiplication. The constant function $f(x) = 0$ is the zero element in $\mathcal{F}$.
(3) The set $\mathbb{R}^{m \times n}$ of all $m \times n$ matrices is a vector space with the usual notion of addition and scalar multiplication of matrices ("entry-by-entry").
(4) The set $\mathbb{R}[x]$ of all polynomials is a vector space with the usual notions of polynomial addition and scalar multiplication. Even though it is possible to multiply polynomials, this multiplication is *not part of the vector space structure* of $\mathbb{R}[x]$.
(5) The set of all solutions $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ to the equation $5x_1 + 3x_2 - x_3 = 0$ is vector space. Notice two solutions to this equation can be added in an obvious way (componentwise, just like in $\mathbb{R}^3$). Also, we can multiply any solution by a scalar multiple to get another solution. You should run through the axioms of a vector space to convince yourself that the solutions to the equation form a vector space. Indeed, the solutions to *any* linear system of equations of the form $A\vec{x} = \vec{0}$ is a vector space.

Section 1.3.

Definition 1.3. A **linear combination** of vectors $\vec{v}_1, \ldots, \vec{v}_n$ is any vector $\vec{v}$ of the form

$$\vec{v} = a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_n\vec{v}_n.$$

In Section 1.3, the book defines linear combinations only in the context that the $\vec{v}_i$ are (column) vectors in coordinate space $\mathbb{R}^n$. However, linear combinations are defined in any context where "addition and scalar multiplication" make sense—that is, in any *vector space.*

Example 1.4.

(1) Every vector $\begin{bmatrix} a \\ b \end{bmatrix}$ in $\mathbb{R}^2$ is a linear combination of the vectors $\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\vec{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Indeed,

$$\begin{bmatrix} a \\ b \end{bmatrix} = a \begin{bmatrix} 1 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \end{bmatrix} = a\vec{e}_1 + b\vec{e}_2.$$

(2) More generally, every vector in $\mathbb{R}^n$ is a linear combination of the standard unit vectors $\vec{e}_1, \ldots, \vec{e}_n$.

(3) Consider the functions $f(x) = sin^2(x)$ and $g(x) = cos^2(x)$. These are vectors in the vector space of all functions $\mathcal{F}$ from $\mathbb{R}$ to $\mathbb{R}$. The constant function $h(x) = 1$ is a linear combination of f and g. Indeed: for all x, we have

$$sin^2(x) + cos^2(x) = 1$$

so $h(x) = f(x) + g(x)$.

Book Concepts you must master. VOCABULARY: system of linear equations, consistent, inconsistent, matrix, augmented matrix, row vector, column vector, elementary row operation, row-reduced echelon form (rref), leading one (or pivot) in rref, rank, free variables, leading variables, linear combination, matrix addition and multiplication.

IMPORTANT SKILLS: Solving linear systems of equations using the techniques described on page 15 of textbook (see example on page 14-15). Using the book theorems 1.3.1, 1.3.3 and 1.3.4 on interpreting whether or not a linear system has no solutions, exactly one solution, or infinitely many solutions. Computing the rank of a matrix and the row reduced echelon form. Being able to determine when some vector is a linear combination of some given vectors. These are not just computational techniques you can forgot after Chapter 1: all the theory later will be built upon these technique.

2. Chapter 2

The main idea is the concept of linear transformation, and matrices as a tool to understand them.

Section 2.1: Linear Transformations.

Definition 2.1. Let V and W be vector spaces. A **linear transformation** is a mapping $V \xrightarrow{T} W$ that satisfies both of the following two conditions:

(1) $T(\vec{x}+\vec{y}) = T(\vec{x}) + T(\vec{y})$ for all vectors $\vec{x}, \vec{y} \in V$; and
(2) $T(k\vec{x}) = kT(\vec{x})$ for all vectors $\vec{x} \in V$ and all scalars k.

The vector space V is called the **source** or **domain** of T, whereas W is the **target** of T.

Example 2.2. Let A be an $n \times m$ matrix. The map

$$\mathbb{R}^m \xrightarrow{T_A} \mathbb{R}^n$$

defined by left multiplication by A

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \mapsto A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$$

is a linear transformation. We call this the *transformation given by left multiplication by the matrix A*.

Proof of Example: To prove T_A is linear, we need to show that for all column vectors $\vec{x}, \vec{y} \in \mathbb{R}^m$ and all scalars k, that

$$T_A(\vec{x}+\vec{y}) = T_A(\vec{x}) + T_A(\vec{y}) \text{ and } T_A(k\vec{x}) = kT_A(\vec{x}),$$

that is, that T_A respects addition and scalar multiplication. This follows from the definition of T_A and basic properties of matrix multiplication:

$$T_A(\vec{x}+\vec{y}) = A(\vec{x}+\vec{y}) = A\vec{x} + A\vec{y} = T_A(\vec{x}) + T_A(\vec{y})$$

(the second equality comes from the distributive property of matrix multiplication) and

$$T_A(k\vec{x}) = Ak\vec{x} = kA\vec{x} = kT_A(\vec{x}).$$

□

Remark 2.2.1. Somewhat misleadingly, the book gives *only* this example of a linear transformation (and calls it the "definition"). However, there are many important examples of vector spaces and linear transformations throughout mathematics, science and engineering, many of which are already familiar to you. In Math 217, you must use the definition of linear transformation in this document and be able to apply it to examples beyond coordinate space.

Example 2.3. Here are a few familiar examples of linear transformations. Identify the SOURCE and TARGET, and verify each is a linear transformation.

(1) The *rotation map* of the Cartesian plane rotating vectors counterclockwise by $\frac{\pi}{2}$;
(2) The *evaluation map* $f(x) \mapsto f(0)$ from the space of all functions to $\mathbb{R}$;
(3) The *differentiation map* on the space $\mathcal{C}^\infty$ of infinitely differentiable functions.
(4) Let $\mathbb{R}^{m\times n}$ be the vector space of $m \times n$ matrices. Fix any $n \times n$ matrix A. Then the map

$$\mathbb{R}^{m\times n} \xrightarrow{R_A} \mathbb{R}^{m\times n} \quad \text{defined by} \quad B \mapsto BA$$

is a linear transformation. Similarly, if C is an $m \times m$ matrix, then the map

$$\mathbb{R}^{m\times n} \xrightarrow{L_C} \mathbb{R}^{m\times n} \quad \text{defined by} \quad B \mapsto CB$$

is a linear transformation. So both left and right multiplication by a fixed matrix is linear. Be sure you can prove this, using the basic properties of matrix multiplication.

IMPORTANT SPECIAL CASE: The **coordinate space** $\mathbb{R}^n$ is the most important case of a vector space because as we will soon see, many vector spaces can be modelled on $\mathbb{R}^n$. There is a useful concrete description of all linear transformations $\mathbb{R}^m \to \mathbb{R}^n$:

Crucial Theorem 2.4. *Let* $\mathbb{R}^m \xrightarrow{T} \mathbb{R}^n$ *be a linear transformation. There exists a* **unique** *matrix* A *such that for all* $\vec{x} \in \mathbb{R}^m$*, we have*

$$T(\vec{x}) = A\vec{x}.$$

Morever, the matrix A *is the* $n \times m$ *matrix*

$$A = \begin{bmatrix} T(\vec{e}_1) & T(\vec{e}_2) & \dots & T(\vec{e}_m) \end{bmatrix},$$

whose columns are the images of the standard unit vectors $\vec{e}_j$ *under* T.

The **Crucial Theorem** guarantees that *every linear transformation*

$$\mathbb{R}^m \xrightarrow{T} \mathbb{R}^n$$

of coordinate spaces is a left multiplication by some matrix A*,* and gives us a recipe to find A. This partially justifies why the book *defines* a linear transformation as a matrix multiplication (see Definition 2.1.1 in the book). However, since many important vector spaces *are not just coordinate space*, the book's definition is inadequate for the more sophisticated mathematical treatment you need to understand in Math 217.

The proof of the Crucial Theorem uses the following **Unreasonably Useful Lemma**:

Lemma 2.5. *Let A be an $n \times m$ matrix. The j-th column of A is the matrix product*

$$A\vec{e}_j$$

where $\vec{e}_j$ is the j-th standard unit column vector.

Proof of Lemma: Exercise. $\square$

Proof of the Crucial Theorem. Take $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$ from the source $\mathbb{R}^m$. Write

$$\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_m\vec{e}_m.$$

Applying T, and using the fact that T is linear, we have

$$\begin{aligned} T(\vec{x}) =& T(x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_m) \\ =& x_1T(\vec{e}_1) + x_2T(\vec{e}_2) + \cdots + x_mT(\vec{e}_m) \\ =& \begin{bmatrix} T(\vec{e}_1) & T(\vec{e}_2) & \dots & T(\vec{e}_m) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}. \end{aligned}$$

Thus for any vector $\vec{x}$, we have $T(\vec{x}) = A\vec{x}$, where A is the matrix as defined in the theorem.

It remains to be shown that A is the unique matrix with this property. Suppose, on the contrary, that there is some other matrix B such that

$$T(\vec{x}) = B\vec{x}$$

for all vectors $\vec{x} \in \mathbb{R}^m$. Then for all vectors $\vec{x} \in \mathbb{R}^m$, we have

$$A\vec{x} = B\vec{x}.$$

In particular, taking $\vec{x}$ to be the standard unit column vector $\vec{e}_j$, we have

$$A\vec{e}_j = B\vec{e}_j$$

for each $\vec{e}_j \in \mathbb{R}^m$. By the Unreasonably Useful Lemma, we conclude that the j-th columns and A and B are the same. Since this holds for each of the n columns of the matrices A and B, we conclude that $A = B$. $\square$

Definition 2.6. An **isomorphism** of vector spaces is a *bijective* (or *invertible*) linear transformation.
Vector spaces V and W are **isomorphic** if there exists an isomorphism $V \xrightarrow{T} W$.
The **inverse** of an isomorphism $V \xrightarrow{T} W$ is the unique map $W \xrightarrow{T^{-1}} V$ assigning to each $\vec{w} \in W$ the unique vector $\vec{v}$ in V such that $T(\vec{v}) = \vec{w}$.

Example 2.7. The *transpose map* sending each row vector $\begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix}$ to the corresponding column vector $\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$ defines an isomorphism

$$\mathbb{R}^{1\times n} \to \mathbb{R}^{n\times 1} = \mathbb{R}^n.$$

Of course, the space of row vectors is not really any different from the space of column vectors, since we can write any row as a column and vice-versa. This intuitive idea of "essentially the same after renaming" is exactly what we mean by *isomorphism.*

Example 2.8. Consider the map $\mathbb{C} \to \mathbb{R}^2$ sending a complex number $x + iy$ to the point $\begin{bmatrix} x \\ y \end{bmatrix}$ in the coordinate plane. This is an isomorphism, as you should check.

Example 2.9. Let A be any invertible $n \times n$ matrix. The map $\mathbb{R}^n \xrightarrow{T_A} \mathbb{R}^n$ given by left multiplication by A is an invertible linear transformation, hence an **isomorphism.** To prove that this map is bijective, we need to show that for every $\vec{y}$ in the target $\mathbb{R}^n$, there is a **unique** $\vec{x}$ in the source $\mathbb{R}^n$ such that $T_A(\vec{x}) = \vec{y}$. This is easy: take $\vec{x} = A^{-1}\vec{y}$ (check it!). The inverse map is given by multiplication by the inverse matrix A^{-1}. There are many different invertible $n \times n$ matrices, hence many different self-isomorphisms[2] of $\mathbb{R}^n$.

Proposition 2.10. *Let $V \xrightarrow{T} W$ be an isomorphism of vector spaces. The inverse map $W \xrightarrow{T^{-1}} V$ is also linear, hence also an isomorphism.*

Proof. We need to check that

$$T^{-1}(\vec{w}_1 + \vec{w}_2) = T^{-1}(\vec{w}_1) + T^{-1}(\vec{w}_2) \quad \text{and} \quad T^{-1}(k\vec{w}_1) = kT^{-1}(\vec{w}_1)$$

for all $\vec{w}_1, \vec{w}_2 \in W$ and all scalars k.

Suppose $T^{-1}(\vec{w}_1) = \vec{v}_1$ and $T^{-1}(\vec{w}_2) = \vec{v}_2$. By definition of T^{-1}, this means that $T(\vec{v}_1) = \vec{w}_1$ and $T(\vec{v}_2) = \vec{w}_2$. So because T is linear, $T(\vec{v}_1 + \vec{v}_2) = T(\vec{v}_1) + T(\vec{v}_2) = \vec{w}_1 + \vec{w}_2$, which shows that $T^{-1}(\vec{w}_1 + \vec{w}_2) = \vec{v}_1 + \vec{v}_2$ (again, by the definition of the inverse map T^{-1}). Thus T^{-1} respects addition.

Likewise, linearity of T guarantees that $T(k\vec{v}_1) = kT(\vec{v}_1) = k\vec{w}_1$. This means that $T^{-1}(k\vec{w}_1) = k\vec{v}_1$. So $T^{-1}(k\vec{w}_1) = kT^{-1}(\vec{w}_1)$, showing that T^{-1} respects scalar multiplication as well. □

We can think of an isomorphism as a "renaming" of elements. The inverse isomorphism T^{-1} "undoes the renaming."

[2] Often called *automorphisms*, with the prefix "auto" meaning "self."

Section 2.3: Composition of linear transformations and matrix products.

Theorem 2.11. *A composition of linear transformations* $V \xrightarrow{T} W \xrightarrow{S} Q$, *where* V, W *and* Q *are vector spaces, is linear.*

Proof. We need to verify that $S \circ T$ satisfies the two conditions of Definition 2.1:

(1) $(S \circ T)(\vec{x} + \vec{y}) = (S \circ T)(\vec{x}) + (S \circ T)(\vec{y})$ for all vectors $\vec{x}, \vec{y}$ in V; and
(2) $(S \circ T)(k\vec{x}) = k(S \circ T)(\vec{x})$ for all vectors $\vec{x} \in V$ and all scalars k.

First,
$(S\circ T)(\vec{x}+\vec{y}) = S(T(\vec{x}+\vec{y})) = S(T(\vec{x})+T(\vec{y})) = S(T(\vec{x}))+S(T(\vec{y})) = (S\circ T)(\vec{x})+(S\circ T)(\vec{y})$,
where the linearity of T justifies the second equality and the linearity of S justifies the third. So $S \circ T$ respects addition.

Similarly,
$$(S \circ T)(k\vec{x}) = S(T(k\vec{x})) = S(kT(\vec{x})) = kS(T\vec{x})) = k(S \circ T)(\vec{x}).$$
So $S \circ T$ respects scalar multiplication. Thus $S \circ T$ is linear. □

What about composing maps of coordinate spaces? In this case, we know that linear maps are given by matrix multiplication. What is the matrix of a composition?

Theorem 2.12. *Consider a composition of linear transformations* $\mathbb{R}^m \xrightarrow{T_A} \mathbb{R}^n \xrightarrow{T_B} \mathbb{R}^p$, *where* A *is the matrix of* T_A *and* B *is the matrix of* T_B. *Then the matrix of the composition* $T_B \circ T_A$ *is* BA. *That is,*
$$T_B \circ T_A = T_{BA}.$$

Proof. We compute
$$(T_B \circ T_A)(\vec{x}) = T_B(A\vec{x}) = B(A\vec{x}) = (BA)\vec{x}$$
where the third equality comes from the associative property of matrix multiplication. So the map $T_B \circ T_A$ is the same as left multiplication by the matrix BA. □

Section 2.4: Invertibility of Linear Transformations and Matrices.

Definition 2.13. A matrix A is **invertible** if there exists a matrix B such that $AB = BA = I_n$, an $n \times n$ identity matrix.

The matrix B is called the **inverse** of A. Of course, A is the inverse of B as well. We write A^{-1} for the inverse of A, when it exists.

Invertible matrices are square (you proved this on Homework Set 2). If you *already know* A *is square*, the next Proposition will save you a lot of effort in checking invertibility:

Proposition 2.14. *Let A be an $n \times n$ matrix. Then A is invertible if and only if there exists a matrix B such that $AB = I_n$ or such that $BA = I_n$. That is, for a square matrix A we only need to check one of the products AB or BA is the identity in order to conclude A is invertible.*

Proof. This is proved in the textbook. See Theorem 2.4.8 there. □

Proposition 2.15. *A linear transformation $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^n$ is invertible if and only if the corresponding matrix is invertible.*

Proof. First assume $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^n$ is invertible. Let A be the $n \times n$ matrix such that $T(\vec{x}) = A\vec{x}$ for all $\vec{x} \in \mathbb{R}^n$. Because T is invertible, its inverse T^{-1} is also linear (by Proposition 2.10). Let B be the matrix of T^{-1}. The composition $T \circ T^{-1}$ is the identity map, hence its matrix is the identity matrix. On the other hand, its matrix is also AB, using Theorem 2.12. Hence $AB = I_n$. Similarly, because $T^{-1} \circ T$ is the identity map, we conclude $BA = I_n$. So $AB = BA = I_n$, and A is invertible.

Conversely, assume that A is invertible, with inverse A^{-1}. To see that the linear map T_A given by multiplication by A is invertible, we observe that the map $T_{A^{-1}}$ given by multiplication by A^{-1} is the inverse of T_A. Indeed, $T_A \circ T_{A^{-1}} = T_{A^{-1}} \circ T_A = T_{I_n}$ by Theorem 2.12, and multiplication by I_n is clearly the identity map. □

Proposition 2.16. *Let A be an $n \times n$ matrix. Thinking of A as the coefficient matrix of a system of n linear equations in n unknowns, we have that A is invertible if and only if the system $A\vec{x} = \vec{b}$ has a unique solution for all $\vec{b} \in \mathbb{R}^n$. Moreover, if there is a unique solution for one $\vec{b} \in \mathbb{R}^n$, there is a unique solution for every $\vec{b} \in \mathbb{R}^n$.*

Proof. This is essentially a rephrasing of Proposition 2.15 in terms of linear equations. Let $\mathbb{R}^n \xrightarrow{T_A} \mathbb{R}^n$ be the linear transformation $\vec{x} \mapsto A\vec{x}$. We know that T_A is invertible (or bijective) if and only if the matrix A is invertible. Furthermore, by definition, T_A is bijective if and only if for all $\vec{b}$ in the target $\mathbb{R}^n$, there is a unique $\vec{x}$ in the source $\mathbb{R}^n$ such that $T_A(\vec{a}) = \vec{b}$. That is, A is invertible if and only if the system $A\vec{x} = \vec{b}$ has a unique solution for all $\vec{b} \in \mathbb{R}^n$.

The final statement is true because A is square: think about the process of row-reducing the augmented matrix $A \,|\, \vec{b}$ to find the solutions. There is a unique solution if and only if $rref(A)$ is the identity matrix, regardless of what $\vec{b}$ is. □

Remark 2.16.1. With notation as in the Proposition 2.16, the unique solution to $A\vec{x} = \vec{b}$ is of course the vector $A^{-1}\vec{b}$. Check it!

Remark 2.16.2. It is important in Proposition 2.14 that A is *square*. The statement *"If $AB = I_n$ then A is invertible"* is false without the hypothesis that the matrices are square. Here is a counterexample:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \text{but} \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Book Concepts you must Master from Chapter 2. VOCABULARY: linear transformation, matrix of a linear transformation, coordinate space $\mathbb{R}^n$, standard unit vectors $\vec{e}_j$, domain (or source), target, image, injective, surjective, bijective, invertible map, invertible matrix, inverse map, inverse matrix, names for basic algebraic properties (commutative, associative, distributive, additive or multiplicative identity, additive or multiplicative inverse, etc).

IMPORTANT SKILLS: Verifying given maps are linear transformations using Definition 2.1, recognizing common linear transformation (including rotations, projections, reflections), finding the matrix of a linear transformation (e.g. by computing the columns, thinking about the images of $\vec{e}_j$). You must be able to use Theorems 2.1.2 and 2.1.3.

You should be fast and accurate at multiplying matrices, and be adept at thinking of matrix multiplication in many ways (eg, in Thm 2.3.2 for AB you could write B as a row of columns $[C_1 \;\; C_2 \;\; \cdots \;\; C_n]$ and then AB is the matrix $[AC_1 \;\; AC_2 \;\; \cdots \;\; AC_n]$). You should know how to find the matrix of a composition of linear transformations. You should have an arsenal of counterexamples ready: matrices A and B that don't commute, non-zero matrices A and B such that $AB = 0$, etc.

You must know how to find the INVERSE of a given matrix. Make sure you can use the technique explained in Theorem 2.4.5 (and demonstrated just prior, starting with Example 1 on page 90 but summarized succintly near the top of page 91). You should be able to immediately write down the inverse of a 2×2 matrix; see Theorem 2.4.9. You should be able to tell if a matrix is invertible (Thms 2.4.3, 2.4.7, 2.4.8, 2.4.9) and understand how inverse matrices come up in solving systems (Thm 2.4.4). You should also know all the equivalent characterizations of invertible matrices in Summary 3.1.8 on page 118.

3. Chapter 3

The main idea is the concept of a **BASIS** for a vector space. This gives us the important notion of **DIMENSION** of vector space.

A Deep idea is that bases allow us introduce **COORDINATES** for any vector space, so we can model any (finite dimensional) vector space on the coordinate space $\mathbb{R}^n$ and any linear transformation by left matrix multiplication.

Section 3.1. Span, Kernel and Image.

Definition 3.1. The **span** of a set $\{\vec{v}_1, \ldots, \vec{v}_n\}$ of vectors is the set *all* linear combinations. That is,

$$Span\{\vec{v}_1, \ldots, \vec{v}_n\} = \{c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n \,|\, c_i \in \mathbb{R}\}.$$

Example 3.2. (1) Let $\vec{e}_1$ and $\vec{e}_2$ be the standard unit vectors in $\mathbb{R}^3$. Their span in $\mathbb{R}^3$ is the xy-plane, or in set-notation, $\{ \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} \,|\, x, y \in \mathbb{R}\}$. Similarly, the vectors

$$\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

also span a plane through the origin in $\mathbb{R}^3$. It is the plane $\{ \begin{bmatrix} x \\ y \\ -x \end{bmatrix} \,|\, x, y \in \mathbb{R}\}$. We can also describe it as the plane in $\mathbb{R}^3$ defined by the equation $x = z$.

(2) Consider the subset $\{1, x, x^2, x^3\}$ of the vector space of all polynomials. Its span is the set of all polynomials of degree at most three.

(3) The most obvious spanning set for the coordinate space $\mathbb{R}^3$ is the set of the three standard unit vectors $\vec{e}_1, \vec{e}_2, \vec{e}_3$. However, this is definitely not the only spanning set. The space $\mathbb{R}^3$ is also spanned by $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$, and $\begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}$, although this is less obvious (prove it!). It is also spanned by the four vectors $\vec{e}_1, \vec{e}_2, \vec{e}_3$, and $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, although one might reasonably argue that the fourth vector is redundant.

(4) A line through the origin in $\mathbb{R}^3$ is a vector space. It is spanned by any non-zero vector in it. Do you see why?

(5) The vector space $\mathbb{R}[x]$ of all polynomials is spanned by the polynomials $\{1, x, x^2, x^3, \ldots\}$. No finite subset spans $\mathbb{R}[x]$. Do you see why?

Remark 3.2.1. We can also define the span of an infinite set $\mathcal{S}$ as the collection of all linear combinations of elements in $\mathcal{S}$. That is, Span $\mathcal{S} = \{c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n \,|\, \vec{v}_i \in \mathcal{S}\ c_i \in \mathbb{R}\}$.

Definition 3.3. The **kernel** of a linear transformation $V \xrightarrow{T} W$ of vector spaces is the set of all vectors $\vec{v}$ in the source such that $T(\vec{v}) = \vec{0}$. That is,

$$\ker T = \{\vec{v} \in V \,|\, T(\vec{v}) = \vec{0}\}.$$

Definition 3.4. The **image** of a linear transformation $V \xrightarrow{T} W$ of vector spaces is the set of all vectors $\vec{w}$ in the target such that there exists $\vec{v}$ in the source such that $T(\vec{v}) = \vec{w}$. That is,

$$\operatorname{im} T = \{\vec{w} \in W \,|\, \text{there exists some } \ \vec{v} \in V \ \text{ with } T(\vec{v}) = \vec{w}\}.$$

Proposition 3.5. *A linear transformation $V \xrightarrow{T} W$ is injective if and only if its kernel is trivial.*

Proof. Suppose T is injective. Take $\vec{v}$ in the kernel of T. Then $T(\vec{v}) = 0$. But also $T(\vec{0}) = \vec{0}$. So $\vec{v}$ and $\vec{0}$ have the same image under T. By definition of injective, $\vec{v} = \vec{0}$. So the kernel of T can contain only the zero element.

Conversely, suppose ker T is zero. If $T(\vec{v}) = T(\vec{w})$, then because T is linear, $T(\vec{v} - \vec{w}) = \vec{0}$. So $\vec{v} - \vec{w}$ is in the kernel, making $\vec{v} = \vec{w}$. Thus T is injective. □

An important special case is when our source and target are coordinate spaces $\mathbb{R}^n$. This is the case from which you should draw your intuition, and the only case the book discusses at this point. The next two propositions describe how to think about the kernel and image in these cases.

Theorem 3.6. *Let $\mathbb{R}^m \xrightarrow{T_A} \mathbb{R}^n$ be defined by $T_A(\vec{x}) = A\vec{x}$ for some $n \times m$ matrix A. Then*

(1) *The kernel of T_A is the space of solutions to the linear system $A\vec{x} = \vec{0}$.*
(2) *The image of T_A is the span of the columns of A.*

Proof. For (1): The kernel of T_A is the set of $\vec{x}$ such that $T_A\vec{x} = \vec{0}$. By definition of T_A, this is the set of all $\vec{x}$ such that $A\vec{x} = \vec{0}$, or the solutions to the linear system $A\vec{x} = \vec{0}$.

For (2): The vectors in the image of T_A are those of the form

$$T_A(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}) = T_A(x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_m\vec{e}_m).$$

Using the linearity of T_A, this is the same as

$$x_1 T_A(\vec{e}_1) + x_2 T_A(\vec{e}_2) + \cdots + x_m T_A(\vec{e}_m).$$

But the vectors $T_A(\vec{e}_j) = A\vec{e}_j$ are the columns of the matrix A (by the Unreasonably Useful Lemma). So the vectors in the image of T_A are those that can be written as linear combinations of the columns of A. That is, the image of T_A is the span of the columns of A. □

Remark 3.6.1. A common abuse of language is to say "kernel A" and "image A" instead of "kernel T_A" and "image T_A." We always interpret the kernel and image "of a matrix" to mean the kernel and image of the corresponding linear transformation given by (left) multiplication by this matrix.

Example 3.7. Consider the map $\mathbb{R}^3 \xrightarrow{T} \mathbb{R}^3$ defined by $T(\begin{bmatrix} x \\ y \\ z \end{bmatrix}) = \begin{bmatrix} x+y+z \\ y \\ x+z \end{bmatrix}$. This map can also be described as $T(\vec{x}) = A\vec{x}$ where $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$. According to the previous proposition, the kernel of T is the solution space of the system of linear equations

$$A\vec{x} = \vec{0},$$

which is the line spanned by $\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$. And the image is the span of the vectors

$$\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

Of course, this image is thus the plane spanned by the two vectors $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, and $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ in $\mathbb{R}^3$, since the third column gives no additional information.

Section 3.2: Subspaces and Bases.

Definition 3.8. A **subspace** of a vector space V is a non-empty subset W which is closed under addition and scalar multiplication. That is, a **subspace** is a non-empty subset W of V such that

(1) The zero vector (of V) is in W.
(2) If $\vec{x}, \vec{y} \in W$, then also $\vec{x} + \vec{y} \in W$;
(3) If $\vec{x} \in W$ and k is any scalar, then also $k\vec{x} \in W$.

In the book, in Chapter 3, the only subspaces considered are subspaces of $\mathbb{R}^n$. The general case comes in Chapter 4.

Example 3.9. Any line or plane *through the origin* in $\mathbb{R}^3$ is a subspace. In fact, these are the *only* subspaces of $\mathbb{R}^3$, besides the zero vector space $\{0\}$ and the whole space $\mathbb{R}^3$.

Every subspace is a vector space in its own right. Thus lines and planes through the origin in $\mathbb{R}^3$ are vector spaces. In reading math, when you see that words "Let V be a vector space," a plane in $\mathbb{R}^3$ or a higher dimensional analog, is a pretty good picture to have in mind. This is the only example the book treats in Chapter 3.

Proof of Example 3.9. A line through the origin consists of all vectors of the form $\{t \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \mid t \in \mathbb{R}\}$. Adding two such we have

$$t_1 \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} + t_2 \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = (t_1 + t_2) \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix},$$

so such a line is closed under addition. Scalar multiplying we have

$$k(t_1 \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}) = (kt_1) \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix},$$

so it is also closed under scalar multiplication.

Similarly, a plane through the origin will consist of all points satisfying the equation $ax + by + cz = 0$. If $\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}$ and $\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}$ are two points in this plane, then $ax_1 + by_1 + cz_1 = 0$ and $ax_2 + by_2 + cz_2 = 0$, so also

$$a(x_1 + x_2) + b(y_1 + y_2) + c(z_1 + z_2) = 0,$$

so $\begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \\ z_1 + z_2 \end{bmatrix}$ lies on on the plane as well. Thus the plane is closed under addition. Likewise, if $ax_1 + by_1 + cz_1 = 0$, then also $a(kx_1) + b(ky_1) + c(kz_1) = 0$. So if $\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}$ is on the plane, then any scalar multiple is also on the plane. □

Remark 3.9.1. Strictly speaking, we do not need to explicitly assume (1) in Definition 3.8 of subspace, because (3) implies (1). [To see this, observe that for any $\vec{x} \in W$, closure under scalar multiplication imples that $0 \cdot \vec{x} = \vec{0} \in W$.[3]] However, sometimes (1) is useful in checking quickly that W is NOT a subspace. The reason we include (1) in the definition above is to emphasize it: experience shows students sometimes forget subspaces always contain $\vec{0}$.

Proposition 3.10. *Let $T : V \to W$ be a linear transformation of vector spaces.*

[3]This argument requires that W has some vector $\vec{x}$ in it to start—that is, that W is non-empty. The empty set is not a vector space.

(1) *The kernel of T is a subspace of the source V.*
(2) *The image of T is a subspace of the target W.*

Proof. (1) To show the kernel of T is a subspace of V, we must show that the ker T is closed under addition and scalar multiplication.

Take $\vec{v}_1$ and $\vec{v}_2$ in ker T. By definition of kernel, $T(\vec{v}_1) = 0$ and $T(\vec{v}_2) = 0$. Because T is linear, we know $T(\vec{v}_1) + T(\vec{v}_2) = T(\vec{v}_1 + \vec{v}_2) = 0$. So $\vec{v}_1 + \vec{v}_2$ is in the kernel of T. This shows the kernel of T is closed under addition.

Take $\vec{v} \in \ker T$. By definition, this means that $T(\vec{v}) = 0$. So for any scalar, $kT(\vec{v}) = k \cdot 0 = 0$. By linearity of T, we have $T(k\vec{v}) = 0$. This shows that $k\vec{v}$ is in the kernel of T. So the kernel is closed also under scalar multiplication. We conclude that the kernel is a subspace.

(2) To show the image of T is a subspace of W, we must show that im T is closed under addition and scalar multiplication.

Take $\vec{w}_1$ and $\vec{w}_2$ in the image of T. By definition, this means $T(\vec{v}_1) = \vec{w}_1$ and $T(\vec{v}_2) = \vec{w}_2$ for some $\vec{v}_1, \vec{v}_2 \in V$. Because T is linear, we have

$$T(\vec{v}_1 + \vec{v}_2) = T(\vec{v}_1) + T(\vec{v}_2) = \vec{w}_1 + \vec{w}_2,$$

so that also $\vec{w}_1 + \vec{w}_2$ is in the image of T. So im T is closed under addition.

Similarly, if $\vec{w}$ is in the image, we can write $T(\vec{v}) = \vec{w}$ for some $\vec{v} \in V$. So for any scalar k,

$$k\vec{w} = kT(\vec{v}) = T(k\vec{v}),$$

showing that $k\vec{w}$ is in the image. So the image of T is closed under both addition and scalar multiplication. We conclude that im T is a subspace. □

Proposition 3.11. *Let V be a vector space and $\mathcal{S}$ any subset. The Span of $\mathcal{S}$ is a subspace of V.*

Proof. We need to show:

(1) If $\vec{v}_1$ and $\vec{v}_2$ are in the span of $\mathcal{S}$, then also $\vec{v}_1 + \vec{v}_2$ is in the span of $\mathcal{S}$;
(2) If $\vec{v}$ is in the span of $\mathcal{S}$ and k is any scalar, then $k\vec{v}$ is in the span of $\mathcal{S}$.

I leave this for you: use the definition of span, write out what it means that a vector in the span of $\mathcal{S}$ and check these two properties. This will complete the proof. *Come see me* if you are not 100 % sure you have done this correctly. □

Linear Independence.

Definition 3.12. A **relation** on a set of vectors $\{\vec{v}_1, \dots, \vec{v}_n\}$ is any expression of the form

$$c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n = 0,$$

where the c_i are scalars. That is, a relation is a linear combination that equals the **zero** vector.

Every set $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ has the **trivial relation:**

$$0 \cdot \vec{v}_1 + 0 \cdot \vec{v}_2 + \cdots + 0 \cdot \vec{v}_n = 0.$$

We say a relation

$$c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n = 0$$

is **non-trivial** if at least one of the coefficients c_i is non-zero.

Definition 3.13. A set of vectors $\{\vec{v}_1, \dots, \vec{v}_n\}$ is **linearly independent** if the only relation is the trivial relation— that is, whenever $c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n = 0$ for some scalars c_i, then $c_1 = c_2 = \cdots = c_n = 0$.

Example 3.14. (1) A one-element set $\{\vec{v}\}$ is linearly independent, provided $\vec{v} \neq 0$.

(2) The vectors $\vec{e}_1, \vec{e}_2, \vec{e}_3$ in $\mathbb{R}^3$ are linearly independent. To prove this, consider a relation

$$c_1\vec{e}_1 + c_2\vec{e}_2 + c_3\vec{e}_3 = 0.$$

Expanding out, this says that $\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \vec{0}$, which is possible only if the c_i are all zero. So every relation on the vectors $\vec{e}_1, \vec{e}_2, \vec{e}_3$ is trivial. By definition, the vectors $\vec{e}_1, \vec{e}_2, \vec{e}_3$ are linearly independent.

Proposition 3.15. *A two-element set $\{\vec{v}_1, \vec{v}_2\}$ is linearly independent if and only if neither vector is a multiple of the other.*

Proof. Suppose that $\{\vec{v}_1, \vec{v}_2\}$ is linearly independent. We use "proof by contradiction" to show that neither vector is a multiple of the other. Suppose on the contrary that $c\vec{v}_1 = \vec{v}_2$. Then we have a relation

$$c\vec{v}_1 + (-1)\vec{v}_2 = 0.$$

Since the relation is non-trivial ($-1 \neq 0$), this contradicts linear independence. A similar argument (reversing the roles of $\vec{v}_1$ and $\vec{v}_2$) shows that if $c\vec{v}_2 = \vec{v}_1$, we also get a contradiction.

Conversely, suppose neither vector is a multiple of the other. We need to show that $\{\vec{v}_1, \vec{v}_2\}$ is linearly independent. If not, then there there is a non-trivial relation

$$c_1\vec{v}_1 + c_2\vec{v}_2 = 0.$$

At least one of the c_i is not zero, say $c_1 \neq 0$. Then rearranging, we have $\vec{v}_1 = \frac{-c_2}{c_1}\vec{v}_2$, contrary to assumption. We can argue similarly if $c_2 \neq 0$. □

Remark 3.15.1. You can also define relations and linear independence for an infinite set $\mathcal{S}$. An (arbitrary) set $\mathcal{S}$ of vectors is linearly independent if $c_1\vec{v}_2 + \cdots + c_n\vec{v}_n = 0$ for some $\vec{v}_i$ vectors in $\mathcal{S}$ and some scalars c_i in $\mathbb{R}$, then $c_1 = c_2 = \cdots = c_n = 0$. That is, a (possibly infinite) set $\mathcal{S}$ is linearly independent it has no non-trivial relations.

The definition of **linear independence** is probably the trickiest so far. Please memorize it as stated here; doing so will help you with proofs. For vectors $\vec{v}_1, \ldots, \vec{v}_m$ in $\mathbb{R}^n$, the book gives many useful ways to think about whether or not they are linearly independent. Please study Summary 3.2.9 carefully on page 129.

Bases.

Definition 3.16. A **basis** for a vector space V is a set of vectors which is BOTH *linearly independent* and *spans* V.

Example 3.17. Some natural bases for familiar vector spaces:

(1) A basis for $\mathbb{R}^n$ is the set $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$. This is called the **standard basis.**
(2) A basis for the space $\mathbb{R}^{2\times 2}$ of 2×2 matrices is

$$\{\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}, \begin{bmatrix}0 & 1\\0 & 0\end{bmatrix}, \begin{bmatrix}0 & 0\\1 & 0\end{bmatrix}, \begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}\}.$$

(3) A basis for the vector space $\mathbb{C}$ of complex numbers is $\{1, i\}$.
(4) A basis for the vector space $\mathcal{P}_4$ of polynomials of degree at most four is

$$\{1, x, x^2, x^3, x^4\}.$$

(5) A basis for the vector space $\mathbb{R}[x]$ of *all* polynomials is $\{1, x, x^2, \ldots\}$.
(6) The plane W in $\mathbb{R}^3$ defined by $x+y+z = 0$ is a vector space with basis $\{\begin{bmatrix}1\\0\\-1\end{bmatrix}, \begin{bmatrix}0\\1\\-1\end{bmatrix}\}$.
This example shows that not every vector space has an obvious or natural basis. For example, another basis for W is $\{\begin{bmatrix}1\\-1\\0\end{bmatrix}, \begin{bmatrix}0\\-1\\1\end{bmatrix}\}$.

Theorem 3.18. *If a subset $\mathcal{B}$ is a basis for a vector space V, then every element $\vec{v}$ in V can be written uniquely as a linear combination vectors in $\mathcal{B}$.*

Proof. Take $\vec{v} \in V$. Because $\mathcal{B}$ spans V, we know that $\vec{v}$ can be written as a linear combination of the elements of $\mathcal{B}$. To show this is unique, suppose we can write $\vec{v}$ as a linear combination in two ways. Write

$$\vec{v} = a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_n\vec{v}_n = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n$$

where the $\vec{v}_i$ are in $\mathcal{B}$ and the a_i and c_i are scalars (some could be zero). Subtract one expression from the other to get

$$0 = (a_1 - c_1)\vec{v}_1 + (a_2 - c_2)\vec{v}_2 + \cdots + (a_n - c_n)\vec{v}_n.$$

Because the $\vec{v}_i$ are elements of a basis, they are linearly independent. So by definition, this relation on the $\vec{v}_i$ must be trivial. This means that $a_i = c_i$ for all i. Thus the expression for $\vec{v}$ as a linear combination of the basis elements is unique. □

The next two theorems are helpful for our intuition about bases:

Theorem 3.19. *Let V be a vector space.*

(1) *A set of vectors is a basis if and only if it is a minimal spanning set.*
(2) *A set of vectors is a basis if and only if it is a maximal linearly independent set.*

We write down the proof of Theorem 3.19 for finite sets of vectors only. The proof is the same for infinite sets but the notation is somewhat more clumsy.

Proof of Theorem 3.19. (1) Let $\{\vec{v}_1, \ldots, \vec{v}_n\}$ be a minimal spanning set of vectors in V. This means that removing any vector from this set will produce a set that fails to span V. To show this set is a basis, we only need to show it is linearly independent, since we already know it spans.

Suppose, on the contrary, that we have a non-trivial relation

$$c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n = 0.$$

Some coefficient, say c_i, is not zero. Rearranging, we have

$$\vec{v}_i = -\frac{c_1}{c_i}\,\vec{v}_1 - \frac{c_2}{c_i}\,\vec{v}_2 - \cdots - \hat{\imath} - \cdots - \frac{c_n}{c_i}\,\vec{v}_n,$$

where the notation $\hat{\imath}$ means the $\vec{v}_i$ term is omitted. This shows that $\vec{v}_i$ is in the span of the smaller set $\{\vec{v}_1, \ldots, \hat{\imath}, \ldots, \vec{v}_n\}$. So the smaller set $\{\vec{v}_1, \ldots, \hat{\imath}, \ldots, \vec{v}_n\}$ in fact spans all of V, contrary to the minimality of the original spanning set. This contradiction establishes that the set $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is linearly independent. So it is a basis.

For the converse, assume $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is a basis. To prove that it is a minimal spanning set, suppose on the contrary, that removing (say, after relabling the vectors) $\vec{v}_n$ is also a spanning set for V. This means that $\vec{v}_n = c_1\vec{v}_1 + \cdots + c_{n-1}\vec{v}_{n-1}$. But then we have the relation $c_1\vec{v}_1 + \cdots + c_{n-1}\vec{v}_{n-1} - \vec{v}_n = 0$, contrary to the assumption that the set is basis.

(2) Let $\{\vec{v}_1, \ldots, \vec{v}_n\}$ be a maximal set of linearly independent vectors in V. This means that adding any vector to this set will make it linearly *dependent*. We only need to check that the set *spans* V, since we know already the set is linearly independent.

Take an arbitrary $\vec{w} \in V$. Since $\{\vec{v}_1, \ldots, \vec{v}_n, \vec{w}\}$ is linearly dependent, there is a non-trivial relation

$$c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n + a\vec{w} = 0.$$

Note that $a \neq 0$, because otherwise we would have a non-trivial relation on the set $\{\vec{v}_1, \dots, \vec{v}_n\}$, contrary to assumption. Hence, rearranging, we have

$$\vec{w} = -\frac{c_1}{a}\,\vec{v}_1 - \frac{c_2}{a}\,\vec{v}_2 - \cdots - \frac{c_n}{a}\,\vec{v}_n.$$

This says that $\vec{w}$ is in the span of $\{\vec{v}_1, \dots, \vec{v}_n\}$. So the set $\{\vec{v}_1, \dots, \vec{v}_n\}$ spans V and hence is a basis.

Conversely, assume $\{\vec{v}_1, \dots, \vec{v}_n\}$ is a basis. Suppose it is not a maximal linearly independent set. Then we can add some vector $\vec{w}$ so that $\{\vec{v}_1, \dots, \vec{v}_n, \vec{w}\}$ is linearly independent. Because $\{\vec{v}_1, \dots, \vec{v}_n\}$ spans V, we can write $\vec{w} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n$ for some scalars c_i. But then $c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n - w = 0$ is a non-trivial relation, contrary to the linear independence of $\{\vec{v}_1, \dots, \vec{v}_n, \vec{w}\}$. This contradiction establishes that $\{\vec{v}_1, \dots, \vec{v}_n\}$ is a maximal linearly independent set. □

Theorem 3.20. *Every vector space has a basis.*

Proof. This is actually a hard theorem to prove in the infinite dimensional case. It is more straightforward in the finite dimensional case. We omit this proof for now. □

Section 3.3. Dimension. Vector spaces typically have many bases, but the number of elements in any basis is always the same:

Theorem 3.21. *All bases of a vector space have the same number (possibly infinite) of elements.*

Proof. Fix a vector space V. If all bases for V are infinite, the theorem holds. So assume V has a finite basis $\{\vec{v}_1, \dots, \vec{v}_n\}$. We need to show every basis for V has n elements. This is proved in the book, Theorem 3.3.2. We will give a different proof in Section 3.4. □

Definition 3.22. The **dimension** of a vector space is the number of vectors in a basis.

Note that the dimension can be infinite.

Example 3.23. Referring to the bases we found in Example 3.17, we see $\mathbb{R}^n$ has dimension n, the space $\mathbb{R}^{2\times 2}$ has dimension 4, the space of complex numbers has dimension 2, the space $\mathcal{P}_4$ (of polynomials of degree four or less) has dimension 5, the space $\mathbb{R}[x]$ (of all polynomials) has infinite dimension, and the plane W has dimension 2.

Theorem 3.24. *Let V be a vector space. The dimension of V is the maximal number of linearly independent vectors in V. Alternatively, the dimension of V is the minimal number of vectors needed to span V.*

Proof. This is really a corollary of Theorem 3.19. Since a maximal set of linearly independent elements in V is a basis (by Theorem 3.19), we know the number of elements in such a set is the dimension. Likewise, since a minimal spanning set is a basis (by Theorem 3.19), we know the number of elements in such a set is the dimension. □

Corollary 3.25. *Let V be a vector space of dimension n and let $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ be any set of n vectors. Then $\mathcal{B}$ spans V if and only if $\mathcal{B}$ is linearly independent.*

Proof. Suppose $\mathcal{B}$ spans V. Since V is n-dimensional, the minimal number of spanning vectors is n by Theorem 3.24. So $\mathcal{B}$ is a minimal spanning set—if we remove any vectors we have a smaller set so it can't span. This means $\mathcal{B}$ is a basis by Theorem 3.19, so the elements are linearly independent.

Conversely, suppose the n vectors of $\mathcal{B}$ are linearly independent. Since V is n- dimensional, this means $\mathcal{B}$ is a maximal linearly independent set—adding any extra vector to it would produce a linearly dependent set. So $\mathcal{B}$ is a basis, and so must also span V. □

Proof Tip: These theorems imply that if V has dimension n **and we have n vectors** $\{\vec{v}_1, \dots, \vec{v}_n\}$, then to check they are a basis we can check **either** they are linearly independent **or** they span V. This can significantly shorten your struggle in many proofs.

The book restates these results in the following useful form:

Book Theorem 3.3.4: Let V be a vector space of dimension m. Then

(1) We can find at most m linearly independent vectors in V.
(2) We need at least m vectors to span V.
(3) Any set of m linearly independent vectors in V is a basis.
(4) Any set of m vectors which spans V is a basis.

Actually, the theorem in the book states this only in the special case that V is a subspace of $\mathbb{R}^n$. The proof is exactly the same.

The Rank-Nullity Theorem. This is my number one most favorite useful theorem of Math 217.

Theorem 3.26. Rank-Nullity Theorem. *Let $V \xrightarrow{T} W$ be a linear transformation, where V is finite dimensional. Then*

$$\dim V = \dim \mathit{ker}\, T + \dim \mathit{im}\, T.$$

That is: "dimension of kernel plus dimension of image = dimension of source."

The reason for the name "Rank-Nullity theorem" comes from some older terminology in linear algebra: The **nullity** of T is the dimension of the kernel. The **rank** of a linear transformation T is the dimension of the image. The next result ensures this terminology is consistent with our previous definition of rank.

Theorem 3.27. *If* $\mathbb{R}^m \xrightarrow{T_A} \mathbb{R}^n$ *is the transformation given by left multiplication by the matrix* A*, the dimension of the image of* T_A *is the rank of the corresponding matrix* A*.*

Proof. This is Theorem 3.3.6 in the book. □

Example 3.28. Consider the linear transformation $\mathbb{R}^m \xrightarrow{T_A} \mathbb{R}^n$ given by left multiplication by the $n \times m$ matrix A. The kernel is the solution space of the linear system

$$A\vec{x} = 0.$$

As you know from chapter one, this solution space will have d free variables, where d is the total number of variables m minus the rank of A (number of leading ones in $rref(A)$). The d free variables means that the kernel is d-dimensional, where $d = m - \text{rank}(A)$. So we recover

$$\dim(source) = \dim(kernel) + \dim(image).$$

Example 3.29. Consider the projection $\mathbb{R}^3 \to \mathbb{R}^2$ sending $\begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} x \\ y \end{bmatrix}$. The kernel is the z-axis, which has basis $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$; hence the kernel has dimension one. The image is all of $\mathbb{R}^2$, hence has dimension two. This confirms that $2 + 1 = 3$, the dimension of the source $\mathbb{R}^3$.

The next theorem is useful in practice to check whether column vectors are linearly independent.

Theorem 3.30. *Column vectors* $\vec{v}_1, \ldots, \vec{v}_d$ *in* $\mathbb{R}^n$ *are linearly independent if and only if the* $n \times d$ *matrix* $[\vec{v}_1 \ \vec{v}_2 \ \ \ldots \vec{v}_d]$ *has rank* d*. In particular, vectors* $\vec{v}_1, \ldots, \vec{v}_n$ *in* $\mathbb{R}^n$ *are a basis for* $\mathbb{R}^n$ *if the* $n \times n$ *matrix formed by its columns is invertible.*

3.4 Coordinates.

Section 3.4 is so important, it should really be its own chapter called Representing Vectors by Columns and Transformations by Matrices. A lot of it will make more sense to you after reading Chapter 4 of the book, especially 4.3. My writeup called "Change of Coordinates and All That" should help, too. This section is really the heart of Math 217–the deepest and hardest and most important. You will need to reread the material many times and in different presentations.[4]

Let V be a vector space with basis $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$. Recall that vector $\vec{v}$ in V can be written uniquely as

$$\vec{v} = a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_n\vec{v}_n$$

for some scalars a_i (by Theorem 3.18).

Definition 3.31. Let V be a vector space with basis $\mathcal{B} = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$. The **$\mathcal{B}$-coordinates** of a vector $\vec{v}$ in V are the unique scalars $a_1, \dots, a_n$ such that

$$\vec{v} = a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_n\vec{v}_n.$$

We usually gather the coordinates into a column vector

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix},$$

called the **$\mathcal{B}$-coordinate column vector** of $\vec{v}$. We also write $[\vec{v}]_{\mathcal{B}}$ for this column vector.

Example 3.32. (1) Let $\mathcal{S}$ be the standard basis $\{\vec{e}_1, \dots, \vec{e}_n\}$ for $\mathbb{R}^n$. Then the $\mathcal{S}$-coordinates of a vector $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ are just the standard coordinates, since

$$\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_n.$$

[4] The book only covers in Chapter 3 a special case of what we do here , though it does get to this in Chapter 4. If you are overwhelmed and just want to make it through Exam 1, it is OK to focus on reading 3.4 from the book for now. For a deeper understanding, you will want to read this as soon as you feel ready, or at the latest, when we are officially covering Section 4.3 in the book after Exam 1. That being said, I think if you feel confident on the material in the earlier parts of this document (at least definitions, example, statements of theorems, and some proofs of more basic theorems), reading this while trying to understand 3.4 in the book will help you understand what is "really going on" so could help on Exam 1 as well. But don't stress over this section instead of practicing computations from the book if that's what you need, since some of this goes beyond what you need for Exam 1. For Exam 1, you only need the material from 3.4 in the book (and of course, everything that came before). That section deals with the special case of Theorem 3.42 when $V = \mathbb{R}^n$ and Theorem 3.53. Both are in the book in Section 3.4.

That is,

$$[\vec{x}]_{\mathcal{S}} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

So the coordinates with respect to the standard basis are simply the usual coordinates you already know.

(2) Let $V \subset \mathbb{R}^3$ be plane spanned by two vectors $\vec{v}_1$ and $\vec{v}_2$, as in Example 1 of page 147 of the textbook. Each point of the plane is a *unique* combination $a\vec{v}_1 + b\vec{v}_2$. The coordinates with respect to the basis $\{\vec{v}_1, \vec{v}_2\}$ are thus $\begin{bmatrix} a \\ b \end{bmatrix}$. Please study this example in the book, which continues through page 148 up to the top of page 149, since I can't draw as nice picture as they have.

(3) Consider the vector space $\mathbb{C}$, with basis $\{1, i\}$. The coordinates of $z = x + yi$ with respect to this basis are $\begin{bmatrix} x \\ y \end{bmatrix}$.

(4) Consider the vector space $\mathbb{R}^{2\times 2}$ of two-by-two matrices with basis

$$\mathcal{B} = \left\{ E_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, E_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}.$$

Then the $\mathcal{B}$-coordinates of the matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ are $\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$.

If we instead use the basis

$$\mathfrak{C} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \right\}$$

for $\mathbb{R}^{2\times 2}$, then the $\mathfrak{C}$-coordinates of a matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ are $\begin{bmatrix} a - c + d \\ b \\ d \\ c - d \end{bmatrix}$.

A CRUCIAL IDEA IS THAT COORDINATES LET US IDENTIFY A VECTOR SPACE WITH $\mathbb{R}^n$:

Theorem 3.33. *Let V be a vector space with basis $\mathcal{B} = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$. The map*

$$V \xrightarrow{L_{\mathcal{B}}} \mathbb{R}^n$$

sending each $\vec{v}$ to its $\mathcal{B}$-coordinates

$$\vec{v} \mapsto [\vec{v}]_{\mathcal{B}}$$

is an isomorphism of vector spaces, called the **coordinate isomorphism** *with respect to $\mathcal{B}$.*

Remember that an *isomorphism* is a *bijective linear transformation*— a way of saying two vector spaces are "essentially the same, just with different names." In this theorem, the target $\mathbb{R}^n$ is *the space of $\mathcal{B}$-coordinates for V*. Basically, the coordinate isomorphism is a "labelling" of all the vectors in V by their coordinates (with respect to the basis $\mathcal{B}$).

Example 3.34. (1) Consider the vector space $\mathbb{C}$, with basis $\{1, i\}$. The "obvious map"

$$\mathbb{C} \to \mathbb{R}^2 \quad \text{sending} \quad x + yi \mapsto \begin{bmatrix} x \\ y \end{bmatrix}$$

is an isomorphism of vector spaces—it is the coordinate isomorphism determined by the basis $\{1, i\}$.

(2) The "obvious map"

$$\mathbb{R}^{2\times 2} \to \mathbb{R}^4 \quad \text{sending} \quad \begin{bmatrix} a & b \\ c & d \end{bmatrix} \mapsto \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$$

is the precisely the coordinate isomorphism induced by the basis $\{E_{11}, E_{12}, E_{21}, E_{22}\}$ described in Example 3.32 above.

(3) If we had picked more exotic bases for either of the example in (1) and (2), we would get different isomorphisms. For example, the basis $\mathfrak{C}$ for $\mathbb{R}^{2\times 2}$ discussed in Example 3.32 gives the isomorphism

$$\mathbb{R}^{2\times 2} \to \mathbb{R}^4 \quad \text{sending} \quad \begin{bmatrix} a & b \\ c & d \end{bmatrix} \mapsto \begin{bmatrix} a - c + d \\ b \\ d \\ c - d \end{bmatrix}.$$

Not all bases are created equal—part of the art of being a good user of linear algebra is choosing convenient bases in which to study your problem.

Example 3.35. Take any basis $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_n\}$ for V. The $\mathcal{B}$ coordinate of the $\vec{v}_i$ are

$$[\vec{v}_1]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad [\vec{v}_2]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \quad [\vec{v}_n]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

So the vector $x_1\vec{v}_1 + \cdots + x_n\vec{v}_n$ is expressed in $\mathcal{B}$-coordinates as

$$[x_1\vec{v}_1 + \cdots + x_n\vec{v}_n]_{\mathcal{B}} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Note that $\vec{v}_j \mapsto \vec{e}_j$ under this isomorphism. So, when we identify V with $\mathbb{R}^n$ using $\mathcal{B}$-coordinates, the basis $\mathcal{B}$ gets identified with the standard basis for $\mathbb{R}^n$.

We can now easily prove that every basis has the same number of elements:

Proof of Theorem 3.21. We have already observed that it suffices to consider only finite bases. Let $\mathcal{B} = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ and $\mathcal{A} = \{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_m\}$ be two different bases for V. We need to show $m = n$. The coordinate isomorphisms

$$V \xrightarrow{L_{\mathcal{B}}} \mathbb{R}^n \quad \text{and} \qquad V \xrightarrow{L_{\mathcal{A}}} \mathbb{R}^m$$

give two different isomorphisms of V with coordinate spaces. The composition

$$\mathbb{R}^n \xrightarrow{L_{\mathcal{B}}^{-1}} V \xrightarrow{L_{\mathcal{A}}} \mathbb{R}^m$$

is an isomorphism as well. The corresponding $m \times n$ matrix of the composition, therefore, is invertible. So $m = n$. □

Remark 3.35.1. Of course, $\mathbb{R}^2$ has many different bases. So there are many different ways to coordinatize $\mathbb{R}^2$ (or indeed any vector space). There is one way you have been studying since middle school—namely the standard Cartesian coordinates (or "x-y coordinates"), which is the coordinate system given by the standard basis. Non-standard coordinates on $\mathbb{R}^n$ can be confusing because we are so brainwashed to think in standard coordinates. You might wonder why one would want to use a *non-standard basis* for $\mathbb{R}^n$. It turns out that for many problems, a clever choice of basis will be very helpful. We will see this already in Examples 3.49 and 3.48 here, but it will be a major theme both in Chapters 5 (where we will choose *orthonormal bases* for $\mathbb{R}^n$) and in Chapter 7 (where we will choose *eigenbases* for $\mathbb{R}^n$).

Remark 3.35.2. Some vector spaces, like $\mathbb{R}^n$, come with a *canonical* (meaning, "natural" or "obvious") choice of basis. For $\mathbb{R}^n$, we can easily argue that the standard basis is a canonical basis. For polynomials, we have the obvious basis $\{1, x, x^2, \ldots\}$. For matrices, the basis as in Example 3.32. Of course, all these vector spaces also have non-standard bases, which depending on the problem you are trying to solve, might turn out to be more convenient.

Remark 3.35.3. Many vector spaces don't come with a "natural choice" of coordinates: think of the plane W defined by $x + y + z = 0$ in $\mathbb{R}^3$. If we fix a basis for W, say

$$\mathcal{B} = \{ \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \},$$

we can "coordinatize" W which allows us to think of this 2-dimensional vector space as a copy of $\mathbb{R}^2$. But an equally reasonable choice of basis is

$$\mathcal{A} = \{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \}.$$

The coordinates of a vector, say $\vec{v} = \begin{bmatrix} 3 \\ 4 \\ -7 \end{bmatrix}$ are different in these bases! Note that

$$\vec{v} = 3 \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + 4 \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix},$$

so

$$[\vec{v}]_{\mathcal{A}} = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \qquad \text{whereas} \qquad [\vec{v}]_{\mathcal{B}} = \begin{bmatrix} -4 \\ 7 \end{bmatrix}.$$

So, if you are a scientist or engineer or mathematician trying to communicate with another, you need to have a systematic way to understand and compare coordinates in different bases.

Comparing different Coordinates. AN IMPORTANT ISSUE TO UNDERSTAND IS THIS: If we have two different bases for V, say $\mathcal{B}$ and $\mathcal{A}$, how do the $\mathcal{B}$-coordinates and $\mathcal{A}$-coordinates compare?

Theorem 3.36. *Let $\mathcal{B}$ and $\mathcal{A}$ be two different bases for an n-dimensional vector space V. The map from the space of $\mathcal{B}$-coordinates to the space of $\mathcal{A}$-coordinates*

$$\mathbb{R}^n \to \mathbb{R}^n \qquad [\vec{v}]_{\mathcal{B}} \mapsto [\vec{v}]_{\mathcal{A}}$$

is a bijective linear map— that is, an isomorphism. In particular, this map is given by left multiplication by the $n \times n$ matrix

$$S_{\mathcal{B}\to\mathcal{A}} = \begin{bmatrix} [\vec{v}_1]_{\mathcal{A}} & [\vec{v}_2]_{\mathcal{A}} & \cdots & [\vec{v}_n]_{\mathcal{A}} \end{bmatrix},$$

whose columns are the elements of the basis $\mathcal{B}$ expressed in $\mathcal{A}$-coordinates.

Definition 3.37. The matrix

$$S_{\mathcal{B}\to\mathcal{A}} = \begin{bmatrix} [\vec{v}_1]_{\mathcal{A}} & [\vec{v}_2]_{\mathcal{A}} & \cdots & [\vec{v}_n]_{\mathcal{A}} \end{bmatrix}.$$

of Theorem 3.36 is called the **change of basis matrix from $\mathcal{B}$ to $\mathcal{A}$.** It can also be defined as the unique matrix such that

$$S_{\mathcal{B}\to\mathcal{A}} \cdot [\vec{v}]_{\mathcal{B}} = [\vec{v}]_{\mathcal{A}}$$

for all vectors $\vec{v}$ in V.

The matrix $S_{\mathcal{B}\to\mathcal{A}}$ transforms the column of $\mathcal{B}$-coordinates of each $\vec{v}$ into its column of $\mathcal{A}$-coordinates, so the change of basis matrix is often called the "change of coordinates matrix."

Example 3.38. Consider the vector space $\mathbb{R}^n$. Let us compare coordinates in the standard basis $\mathcal{S} = \{\vec{e}_1, \dots, \vec{e}_n\}$ to coordinates in some non-standard basis $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ for $\mathbb{R}^n$. What is the matrix that will change the $\mathcal{B}$-coordinates of a vector into the standard coordinates? To figure this out, we need to see where the standard unit vectors $\vec{e}_i$ are taken. We have

$$\vec{e}_i = [\vec{v}_i]_{\mathcal{B}} \mapsto [\vec{v}_i]_{\mathcal{S}}$$

which is simply the column vector $\vec{v}_i$. Hence we have again verified that the change of coordinate matrix is the matrix

$$S_{\mathcal{B}\to\mathcal{A}} = [\vec{v}_1 \;\; \vec{v}_2 \;\; \dots \;\; \vec{v}_n].$$

This is the only case the book considers in Chapter 3. The general case is in Chapter 4.

Caution: Do not confuse $S_{\mathcal{B}\to\mathcal{A}}$ with $S_{\mathcal{A}\to\mathcal{B}}$, which is the matrix transforming $\mathcal{A}$-coordinates into $\mathcal{B}$-coordinates! Of course, these transformations are inverse to each other (do you see why?) Thus their matrices are as well. We state and prove this formally in Proposition 3.41 below.

Example 3.39. The vector space $\mathbb{R}^2$ has nonstandard basis $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix} \right\}$. How do we convert coordinates in this basis to coordinates in the standard basis $\mathcal{A} = \{\vec{e}_1, \vec{e}_2\}$—that is, to standard coordinates?

Theorem 3.36 tells us we must multiply the column of $\mathcal{B}$-coordinates by

$$S_{\mathcal{B}\to\mathcal{A}} = \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}$$

to transform to standard coordinates. Let's check this for the vector $\vec{v} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$. Since $\vec{v} = \vec{v}_1 + \vec{v}_2$, its $\mathcal{B}$-coordinates are $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. We convert to standard coordinates by multiplying by $S_{\mathcal{B}\to\mathcal{A}}$:

$$[\vec{v}]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} = [\vec{v}]_{\mathcal{A}},$$

where is of course $\vec{v}$ expressed in standard coordinates!

Computation Tip: It is almost always easier to find the change of basis matrix **from** a less standard basis **to** a more standard one. Go ahead and directly try to compute the matrix $S_{\mathcal{A}\to\mathcal{B}}$ in Example 3.39 by writing the vectors $\vec{e}_i$ as a linear combination of the $\{\vec{v}_1, \vec{v}_2\}$. You will see what I mean. By contrast, notice how simple it was to find the change of basis matrix $S_{\mathcal{B}\to\mathcal{A}}$ **to** the standard basis.

Example 3.40. Let $\mathcal{P}_1$ be the vector space of polynomials of degree 1 or less. Its elements are the functions[5] of the form $f(x) = mx + b$. It has an "obvious" basis $\mathcal{A} = \{1, x\}$. Another basis is $\mathcal{B} = \{1, x - 1\}$. The change of basis matrix **from** $\mathcal{B}$ to $\mathcal{A}$ is easy to find:

$$S_{\mathcal{B}\to\mathcal{A}} = [[1]_{\mathcal{A}} \;\; [x-1]_{\mathcal{A}}] = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}.$$

We can use this to convert from $\mathcal{B}$-coordinates to $\mathcal{A}$ as follows:

$$[c + d(x-1)]_{\mathcal{B}} = \begin{bmatrix} c \\ d \end{bmatrix} \mapsto \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} c - d \\ d \end{bmatrix} = [c + d(x-1)]_{\mathcal{A}}.$$

This reflects the fact that we can write the polynomial $c + d(x-1)$ as $(c-d) + dx$. To find the matrix $S_{\mathcal{A}\to\mathcal{B}}$, instead of directly computing from Definition 3.37, we can just invert $S_{\mathcal{B}\to\mathcal{A}}$. So

$$S_{\mathcal{A}\to\mathcal{B}} = [S_{\mathcal{B}\to\mathcal{A}}]^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

[5]which are usually called "linear functions" though they are not linear transformations if $b \neq 0$.

Proof of Theorem 3.36. This map is the composition of the coordinate isomorphisms from Theorem 3.33:

$$\mathbb{R}^n \xrightarrow{L_{\mathcal{B}}^{-1}} V \xrightarrow{L_{\mathcal{A}}} \mathbb{R}^n \qquad [\vec{v}]_{\mathcal{B}} \mapsto \vec{v} \mapsto [\vec{v}]_{\mathcal{A}},$$

hence it is also an isomorphism.

Because it is a linear transformation $\mathbb{R}^n \to \mathbb{R}^n$, we know it is given by multiplication by *some* matrix S. We can find S by finding each of its columns. We know that the j-th column of S is the image of $\vec{e}_j$ under the transformation.[6] So let's follow $\vec{e}_j$ through the composition:

$$\mathbb{R}^n \xrightarrow{L_{\mathcal{B}}^{-1}} V \xrightarrow{L_{\mathcal{A}}} \mathbb{R}^n$$
$$\vec{e}_j \mapsto \vec{v}_j \mapsto [\vec{v}_j]_{\mathcal{A}}.$$

Note here that $\vec{e}_j$ maps to $\vec{v}_j$ since $[\vec{v}_j]_{\mathcal{B}} = \vec{e}_j$. Thus the j-th column of the matrix S is $[\vec{v}_j]_{\mathcal{A}}$, as claimed.

□

Proposition 3.41. *With notation as in Definition 3.37, we have* $S_{\mathcal{A}\to\mathcal{B}} = [S_{\mathcal{B}\to\mathcal{A}}]^{-1}$.

Proof. For any vector $\vec{v} \in V$, we have matrix multiplications

$$S_{\mathcal{A}\to\mathcal{B}} \cdot S_{\mathcal{B}\to\mathcal{A}} \cdot [\vec{v}]_{\mathcal{B}} = S_{\mathcal{A}\to\mathcal{B}} \cdot [\vec{v}]_{\mathcal{A}} = [\vec{v}]_{\mathcal{B}},$$

by the definition of the change of basis matrix. So

$$(S_{\mathcal{A}\to\mathcal{B}} \cdot S_{\mathcal{B}\to\mathcal{A}}) \cdot [\vec{v}]_{\mathcal{B}} = [\vec{v}]_{\mathcal{B}},$$

which means that $S_{\mathcal{A}\to\mathcal{B}} \cdot S_{\mathcal{B}\to\mathcal{A}}$ represents the identity map, hence must be I_n. So

$$S_{\mathcal{A}\to\mathcal{B}} \cdot S_{\mathcal{B}\to\mathcal{A}} = I_n.$$

Since both matrices are $n \times n$, we conclude that they are inverse to each other (Proposition 2.14). That is, $S_{\mathcal{A}\to\mathcal{B}} = S_{\mathcal{B}\to\mathcal{A}}^{-1}$. □

Modelling linear transformations by matrix multiplication. Suppose we have a linear transformation

$$V \xrightarrow{T} V.$$

Fix a basis for V, say $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_n\}$. Then we can think of V as "modelled on" $\mathbb{R}^n$ by identifying each vector with its coordinate column. Does this mean we can think of the linear transformation T as "modelled on" a linear transformation

$$\mathbb{R}^n \longrightarrow \mathbb{R}^n$$

of the $\mathcal{B}$-coordinate space? The answer is YES!

[6]These three sentences are all coming from the Crucial Theorem 2.4. If those lines don't make sense, you should go reread the Section 2.1, and talk about it with as many people as you can.

Crucial Theorem 3.42. New and Improved Version. *Let $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_n\}$ be a basis for the vector space V. Let $V \xrightarrow{T} V$ be a linear transformation. Then the corresponding map of $\mathcal{B}$-coordinate columns*

$$[\vec{v}]_{\mathcal{B}} \mapsto [T(\vec{v})]_{\mathcal{B}}$$

is a linear transformation $\mathbb{R}^n \to \mathbb{R}^n$. Moreover, the matrix of this transformation is the $n \times n$ matrix

$$[T]_{\mathcal{B}} = \begin{bmatrix}[T(\vec{v}_1)]_{\mathcal{B}} & [T(\vec{v}_2)]_{\mathcal{B}} & \ldots & [T(\vec{v}_n)]_{\mathcal{B}}\end{bmatrix},$$

whose j-th column is $T(\vec{v}_j)$ expressed in the basis $\mathcal{B}$. That is,

$$[T(\vec{v})]_{\mathcal{B}} = [T]_{\mathcal{B}} \cdot [\vec{v}]_{\mathcal{B}},$$

for every vector $\vec{v} \in V$.

Definition 3.43. The matrix $[T]_{\mathcal{B}}$ in Theorem 3.42 is called the **matrix of T with respect to the basis** $\mathcal{B}$, or simply the **$\mathcal{B}$-matrix** of T. That is, the $\mathcal{B}$ matrix of T is the $n \times n$ matrix

$$[T]_{\mathcal{B}} = \begin{bmatrix}[T(\vec{v}_1)]_{\mathcal{B}} & [T(\vec{v}_2)]_{\mathcal{B}} & \ldots & [T(\vec{v}_n)]_{\mathcal{B}}\end{bmatrix}.$$

Example 3.44. Consider the map $T : \mathbb{C} \to \mathbb{C}$ sending $z \mapsto iz$. It is easy to check that this is linear. The standard identification of $\mathbb{C}$ with $\mathbb{R}^2$ is the coordinate isomorphism defined by the basis $\{1, i\}$ for $\mathbb{C}$. This map identifies $z = x + iy$ with the column vector $\begin{bmatrix} x \\ y \end{bmatrix}$. Note that $T(x + iy) = -y + ix$, so the corresponding linear map of the coordinate-spaces is

$$\mathbb{R}^2 \to \mathbb{R}^2 \quad \text{sending} \quad \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} -y \\ x \end{bmatrix}$$

which has matrix $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. Thus the matrix of T with respect to the basis $\{1, i\}$ is $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. You should also check that its columns are the images of the basis elements 1 and i under T, expressed again in $\mathcal{B}$-coordinates.

Proof of Theorem 3.42. The map $[\vec{v}]_{\mathcal{B}} \mapsto [T(\vec{v})]_{\mathcal{B}}$ is the composition

$$\mathbb{R}^n \xrightarrow{L_{\mathcal{B}}^{-1}} V \xrightarrow{T} V \xrightarrow{L_{\mathcal{B}}} \mathbb{R}^n.$$

Since a composition of linear transformations is linear, this map is linear.

To find the matrix of the composition, we use the Crucial Theorem 2.4. We know the map is given by some matrix, and we just need to figure out which one. We find it by finding each column: we know the j-th column should be the image of $\vec{e}_j$ under this composition. To figure out the image of $\vec{e}_j$ for each j, follow $\vec{e}_j$ through the composition map:

$$\vec{e}_j \mapsto \vec{v}_j \mapsto T(\vec{v}_j) \mapsto [T(\vec{v}_j)]_{\mathcal{B}},$$

where the last vector is the column of $\mathcal{B}$-coordinates for $T(\vec{v}_j)$. [Here, the first arrow $\vec{e}_j \mapsto \vec{v}_j$ is because $[\vec{v}_j]_{\mathcal{B}} = \vec{e}_j$.] So the j-th column of $[T]_{\mathcal{B}}$ is $[T(\vec{v}_j)]_{\mathcal{B}}$, as claimed. □

Theorem 3.42 says that any linear transformation $V \longrightarrow V$ can be treated like a linear transformation $\mathbb{R}^n \to \mathbb{R}^n$ simply by identifying V with the $\mathcal{B}$-coordinate space $\mathbb{R}^n$. Thus, we have a way of thinking of any linear transformation as a matrix multiplication!

Example 3.45. Let $\mathcal{P}_4$ be the vector space of polynomials of degree four or less. Consider the map $d : \mathcal{P}_4 \to \mathcal{P}_4$ sending $f \mapsto f'$. A basis for $\mathcal{P}_4$ is $\mathcal{B} = \{1, x, x^2, x^3, x^4\}$. Under d, we compute the image of each basis element, expressed in $\mathcal{B}$-coordinates, to find the matrix. For example, we compute the third column of the $\mathcal{B}$-matrix in detail as follows:

The third element in the $\mathcal{B}$-basis is x^2.
We apply the transformation d to obtain $2x$.
We then rewrite this result as a linear combination of the basis $\mathcal{B}$ to find the $\mathcal{B}$-coordinates: $0 \cdot 1 + 2 \cdot x + 0 \cdot x^2 + 0 \cdot x^3 + 0 \cdot x^4$.

Thus the third column of the $\mathcal{B}$-matrix is $[d(x^2)]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 2 \\ 0 \\ 0 \\ 0 \end{bmatrix}$.

Doing this for each column, we see that the matrix of d with respect to $\mathcal{B}$ is

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Example 3.46. Let $\mathbb{R}^n \xrightarrow{T_A} \mathbb{R}^m$ be left multiplication by A. What is the matrix of T_A in the standard basis? Using Definition 3.43, we recover precisely A, as you should check! So you have already mastered the process of finding $\mathcal{B}$-matrices in the important case where $V = \mathbb{R}^n$ and $\mathcal{B}$ is the standard basis.

One more time, we rephrase the **Crucial Theorem** and reprove it once more:

Crucial Theorem 3.47. *Let V be a vector space of dimension n with basis $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_n\}$. Let $T : V \to V$ be a linear transformation. Let*

$$[T]_{\mathcal{B}} = \begin{bmatrix} [T(\vec{v}_1)]_{\mathcal{B}} & [T(\vec{v}_2)]_{\mathcal{B}} & \ldots & [T(\vec{v}_n)]_{\mathcal{B}} \end{bmatrix}.$$

be the $\mathcal{B}$-matrix of T. Then for any vector $\vec{v}$, we can compute

"the $\mathcal{B}$-coordinate column vector of $T(\vec{v}) =$ the matrix product $[T]_{\mathcal{B}} \cdot [\vec{v}]_{\mathcal{B}}$."

Proof. Since $\vec{v}_1, \ldots, \vec{v}_n$ is a basis for V, we can write an arbitrary $\vec{v} \in V$ as

$$\vec{v} = x_1\vec{v}_1 + x_2\vec{v}_2 + \ldots x_n\vec{v}_n.$$

So by linearity,

$$T(\vec{v}) = x_1 T(\vec{v}_1) + x_2 T(\vec{v}_2) + \ldots x_n T(\vec{v}_n).$$

So also for coordinate columns:

$$[T(\vec{v})]_{\mathcal{B}} = x_1[T(\vec{v}_1)]_{\mathcal{B}} + x_2[T(\vec{v}_2)]_{\mathcal{B}} + \cdots + x_n[T(\vec{v}_n)]_{\mathcal{B}}.$$

That is

$$[T(\vec{v})]_{\mathcal{B}} = [[T(\vec{v}_1)]_{\mathcal{B}} \ [T(\vec{v}_2)]_{\mathcal{B}} \ \ \ldots \ \ [T(\vec{v}_n)]_{\mathcal{B}}] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

In more compact notation, this says

$$[T(\vec{v})]_{\mathcal{B}} = [T]_{\mathcal{B}} \cdot [\vec{v}]_{\mathcal{B}}$$

for every vector $\vec{v} \in V$. □

Example 3.48. Consider the map $\pi : \mathbb{R}^2 \to \mathbb{R}^2$ given by projection onto the line spanned by $\vec{u}_1$. Let $\vec{u}_2$ be any vector perpendicular to $\vec{u}_1$. Note that $\mathcal{B} = \{\vec{u}_1, \vec{u}_2\}$ is a basis for $\mathbb{R}^2$. What is the $\mathcal{B}$-matrix of π? We compute

$$\pi(\vec{u}_1) = \vec{u}_1 = 1\vec{u}_1 + 0\vec{u}_2, \quad \pi(\vec{u}_2) = 0 = 0\vec{u}_1 + 0\vec{u}_2.$$

So the coordinates of the images of the basis elements are $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$. So the $\mathcal{B}$-matrix is

$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

As you might imagine, the linear transformation in these coordinates is easier to understand than the one we worked out in Chapter 2, Section 2.

Example 3.49. Consider the map $\mathbb{R}^2 \xrightarrow{T} \mathbb{R}^2$ given by multiplication by $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$. How can we understand it geometrically? If we instead use the basis $\mathcal{B} = \{\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}\}$, we compute

$$[T]_{\mathcal{B}} = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}.$$

This means that T stretches all vectors in the $\vec{v}_1$ direction by 3, while fixing all vectors in the direction of $\vec{v}_2$. Sketch the square in the source $\mathbb{R}^2$ determined by the basis $\mathcal{B}$ and its image in the target $\mathcal{B}$.

Comparing matrices of T in Different Bases. An important issue to understand is this: If I have two different bases for V, say $\mathcal{B}$ and $\mathcal{A}$, how do the $\mathcal{B}$-matrix and $\mathcal{A}$ matrix of a linear transformation $V \xrightarrow{T} V$ compare?

Theorem 3.50. *Let V be an n-dimensional vector space, and let $V \xrightarrow{T} V$ be a linear transformation. Suppose that $\mathcal{B}$ and $\mathcal{A}$ are two different bases for V. Then*

$$[T]_{\mathcal{B}} = S^{-1} \cdot [T]_{\mathcal{A}} \cdot S$$

where $S = S_{\mathcal{B}\to\mathcal{A}}$ is the change of basis matrix from $\mathcal{B}$ to $\mathcal{A}$ coordinates.

Definition 3.51. Two $n \times n$ matrices A and B are similar if there exists an invertible $n \times n$ matrix S such that $B = S^{-1}AS$.

Proposition 3.52. *Fix a linear transformation $V \xrightarrow{T} V$. The $\mathcal{B}$-matrices of T in different bases are all similar to each other.*

Proof. If B is the matrix of T with respect to $\mathcal{B}$ and A is the matrix of T with respect to $\mathcal{A}$, then $B = S^{-1}AS$ where $S = S_{\mathcal{B}\to\mathcal{A}}$ is the change of basis matrix from $\mathcal{B}$ to $\mathcal{A}$. □

Proof of Theorem 3.50. We need to check that the two matrices $[T]_{\mathcal{B}}$ and $S^{-1} \cdot [T]_{\mathcal{A}} \cdot S$ are the same matrix. To do this, we can check that they have the same jth column for each $j = 1, 2, \ldots n$. Using the Unreasonably Useful Lemma, we can get at the j-th column of each by multiplying by $\vec{e}_j$.

To compute $S^{-1} \cdot [T]_{\mathcal{A}} \cdot S$, recall that the j-th column of S is $\vec{v}_j$ expressed in the basis $\mathcal{A}$. So

$$(S^{-1} \cdot [T]_{\mathcal{A}} \cdot S) \cdot \vec{e}_j = (S^{-1} \cdot [T]_{\mathcal{A}}) \cdot (S \cdot \vec{e}_j) = (S^{-1} \cdot [T]_{\mathcal{A}}) \cdot [\vec{v}_j]_{\mathcal{A}} = S^{-1} \cdot ([T]_{\mathcal{A}} \cdot [\vec{v}_j]_{\mathcal{A}}).$$

But by definition of $[T]_{\mathcal{A}}$, we have $[T]_{\mathcal{A}} \cdot [\vec{v}]_{\mathcal{A}} = [T(\vec{v})]_{\mathcal{A}}$ for all vectors $\vec{v}$, so in particular,

$$(S^{-1} \cdot [T]_{\mathcal{A}}S)\vec{e}_j = S^{-1} \cdot ([T]_{\mathcal{A}} \cdot [\vec{v}_j]_{\mathcal{A}}) = S^{-1} \cdot [T(\vec{v})]_{\mathcal{A}}.$$

But of course S^{-1} is the matrix which transforms $\mathcal{A}$-coordinates into $\mathcal{B}$-coordinates, so this is

$$[T(\vec{v}_j)]_{\mathcal{B}}.$$

We have just shown that the j-th column of $S^{-1} \cdot [T]_{\mathcal{A}} \cdot S$ is precisely $[T(\vec{v}_j)]_{\mathcal{B}}$. So we have an equality of matrices

$$S^{-1} \cdot [T]_{\mathcal{A}} \cdot S = [T]_{\mathcal{B}}.$$

□

Non-Standard Coordinates on $\mathbb{R}^n$. Of course, $\mathbb{R}^n$ has many basis. How does the $\mathcal{B}$-matrix of a fixed transformation compare to the standard matrix? The preceding discussion specializes as follows:

Theorem 3.53. *Fix a linear transformation $\mathbb{R}^n \xrightarrow{T_A} \mathbb{R}^n$ given by multiplication by the $n \times n$ matrix A. Let $\mathcal{B} = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ be a basis for $\mathbb{R}^n$. Then the $\mathcal{B}$ matrix of T and the standard matrix of T are related by*

$$[T]_{\mathcal{B}} = S^{-1}AS,$$

where S is the matrix formed from the basis $\mathcal{B}$, that is

$$S = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \ldots & \vec{v}_n \end{bmatrix}.$$

This Theorem is really just a special case of Theorem 3.50. Do you see why? The book treats all of the preceding material in Chapter 3 only in the case $V = \mathbb{R}^n$ and compares $\mathcal{B}$-matrices only to standard matrices. They treat the general case in Chapter 4. I personally think it is actually *easier* to understand the general case, where we are less “brainwashed" to rely on standard coordinates.

Book Concepts you must know from Chapter 3. ADDITIONAL VOCABULARY FROM BOOK CHAPTER 3: Image, span, subspace, relation, trivial relation, linearly dependent, linearly independent, basis, dimension, kernel, Rank-Nullity Theorem, coordinates, the change of basis/coordinates matrix, the $\mathcal{B}$-matrix of linear transformation (or the matrix of a linear transformation with respect to basis $\mathcal{B}$), standard matrix, similar matrices.

IMPORTANT SKILLS FROM BOOK CHAPTER 3: checking if vectors are linearly independent, in particular, you should know the characterizations in Summary 3.2.9 of the book. Finding a basis for a vector space, finding the kernel and image of a linear transformation (Example 1 on p 136—be sure to know Theorem 3.3.5 from the book, as well as Theorem 3.3.8), computing the dimension of a vector space, using the rank-nullity theorem, recognizing invertible matrices (eg the summary on page 142), finding coordinates of a vector in a given basis, finding change of coordinate matrices, finding the matrix of a linear transformation in a given basis, converting the matrices of a linear transformation from a representation in one basis to another.